

<https://helda.helsinki.fi>

---

## Different species trait groups of stream diatoms show divergent responses to spatial and environmental factors in a subarctic drainage basin

Lindholm, Marja

2018-07

---

Lindholm , M , Gronroos , M , Hjort , J , Karjalainen , S M , Tokola , L & Heino , J 2018 , ' Different species trait groups of stream diatoms show divergent responses to spatial and environmental factors in a subarctic drainage basin ' , *Hydrobiologia* , vol. 816 , no. 1 , pp. 213-230 . <https://doi.org/10.1007/s10750-018-3585-0>

---

<http://hdl.handle.net/10138/310507>

<https://doi.org/10.1007/s10750-018-3585-0>

---

unspecified

acceptedVersion

---

*Downloaded from Helda, University of Helsinki institutional repository.*

*This is an electronic reprint of the original article.*

*This reprint may differ from the original in pagination and typographic detail.*

*Please cite the original version.*

**Different species trait groups of stream diatoms show divergent responses to spatial and environmental factors in a subarctic drainage basin**

Marja Lindholm<sup>1</sup>, Mira Grönroos<sup>2</sup>, Jan Hjort<sup>1</sup>, Satu Maaria Karjalainen<sup>3</sup>, Laura Tokola<sup>1</sup> and Jani Heino<sup>4</sup>

<sup>1</sup> University of Oulu, Geography Research Unit, P.O. Box 3000, FI-90014 University of Oulu, Finland

<sup>2</sup> University of Helsinki, Department of Environmental Sciences, Section of Environmental Ecology, Niemenkatu 73, FI-15140 Lahti, Finland

<sup>3</sup> Finnish Environment Institute, Freshwater Centre, P.O. Box 413, FI-90014 Oulu, Finland

<sup>4</sup> Finnish Environment Institute, Biodiversity, P.O. Box 413, FI-90014 Oulu, Finland

Correspondence: Marja Lindholm, e-mail address: marja.lindholm@oulu.fi, tel. +358 50 534 7141

**Acknowledgments**

We would like to thank Olli-Matti Kärnä and Sirkku Lehtinen for help in the field and the laboratory. In addition, we would like to thank Annika Vilmi for help in diatom identification and Olli-Matti Kärnä for combining data for Figure 1. This study is part of the project ‘Spatial scaling, metacommunity structure and patterns in stream communities’ funded by the Academy of Finland (project number 273557). During this research, J. Hjort acknowledges the Academy of Finland (project number 285040).

## **Abstract**

Understanding the drivers of community structure is an important topic in ecology. We examined whether different species trait groups of stream diatoms (ecological guilds and specialization groups) show divergent responses to spatial and environmental factors in a subarctic drainage basin. We used local- and catchment-scale environmental and spatial variables in redundancy analysis and variation partitioning to examine community structuring. Local and catchment conditions and spatial variables affected diatom community structure with different relative importance. Local-scale environmental variables explained most of the variation in the low-profile and motile guilds, whereas local and spatial variables explained the same amount of the variation in the high-profile guild. The variations in the planktic guild and the specialist species were best explained by spatial variables, and catchment variables explained most variation only in generalist species. Our study showed that diatom communities in subarctic streams are a result of both environmental filtering and spatial processes. Our findings also suggested that dividing whole community into different groups by species traits can increase understanding of metacommunity organization.

**Keywords:** ecological guilds, ecological specialization, environmental filtering, spatial processes, metacommunity

## Introduction

Understanding the drivers that shape community structure is a central theme in community ecology. These drivers can be studied in the context of a metacommunity (Leibold et al., 2004). A metacommunity is ‘a set of local communities that are linked by dispersal of multiple potentially interacting species’ (Wilson, 1992; Leibold et al., 2004). The concept of metacommunity is based on the notion that the variation in community structure is affected by both local-scale and large-scale environmental and spatial processes (Leibold et al., 2004; Holyoak et al., 2005). It has also been recognized that environmental filtering and dispersal are the fundamental processes structuring metacommunities (Lindström & Langenheder, 2012), as are also biological interactions (Cadotte & Tucker, 2017). Thus, metacommunity studies should focus on the relative roles of these processes (Heino et al., 2015).

The metacommunity has often been treated as a whole without any systematic division within different organismal groups (e.g. diatoms, macrophytes and macroinvertebrates). However, there is typically variation in biological and ecological characteristics between different species even if they belong to the same organismal group (Pandit et al., 2009). The effects of environmental and dispersal processes on local communities may depend on the differences in species traits in metacommunities. Thus, dividing data matrices into different groups by species traits can increase understanding of metacommunity organization (Lindström & Langenheder, 2012). This deconstructive approach has been increasingly applied in recent years when studying community patterns (Grönroos et al., 2013; Alahuhta et al., 2014; Algarte et al., 2014; Vilmi et al., 2017). One way to approach this is to split biological data matrices into smaller parts by dividing species into generalists and specialists based on species ecological specialization (Devictor et al., 2008; Pandit et al., 2009). For example, some studies have shown that environmental control is more dominant in specialist species while generalist respond mainly to spatial processes (e.g. Pandit et al., 2009), whereas other studies have shown different patterns, such as environmental control being dominant independent of specialization (e.g. Székely & Langenheder, 2014). Furthermore, several studies have produced divergent results regarding which factors are important in determining variation in community structure. According to Pandit et al. (2009), these divergent results can be due to different ratios of ecological specialization in different systems studied.

In addition to ecological specialization, biological data matrices can be divided into smaller parts using other biological traits, for example, growth forms and cell sizes (Heino & Soininen, 2006; Rimet & Bouchez, 2012). In the study of freshwater algae, one approach is the use of different guild divisions (Göthe et al., 2013;

Vilmi et al., 2017). Many of these studies have used guild classification based on Passy's (2007) study. Originally, Passy (2007) proposed a diatom guild classification based on the potential of species to use nutrient resources and to resist physical perturbation. Rimet & Bouchez (2012) modified the classification and added one new guild corresponding to planktic species.

Different ecological guilds can be expected to respond in different ways to environmental and spatial processes. Several studies have shown that these guilds respond in different ways to environmental conditions both in lotic (Passy, 2007; Berthon et al., 2011; Rimet & Bouchez, 2012; Göthe et al., 2013) and lentic (Gottschalk & Kahlert, 2012; Vilmi et al., 2017) environments. However, the patterns found have not always been similar, as same guilds have shown dissimilar responses to environment in different studies. Also, these studies have been conducted mainly in areas with relatively high nutrient concentrations, and there is a lack of studies in nutrient poor, harsh subarctic stream environments (but see, Berthon et al., 2011).

In the freshwater realm, studying the relative roles of the environmental and spatial components in community composition is a commonly used approach for understanding metacommunity organization (De Bie et al., 2012; Alahuhta et al., 2014; Vilmi et al., 2016; Vilmi et al., 2017). The environmental components of community variation can be seen as illustrating environmental filtering and the importance of spatial variables may suggest dispersal as determinants of metacommunity structuring (Hájek et al., 2011). Since it is challenging to measure dispersal rates directly (Jacobson & Peres-Neto, 2010), spatial-based dispersal proxies are commonly used (e.g. Grönroos et al., 2013). Specifically, there is very little information available on the dispersal rates of diatom species, and it is particularly difficult to determine dispersal rates of these passively dispersing species directly.

Environmental filtering has been shown to be the main mechanism structuring metacommunities of various organisms in different environments (Van der Gucht et al., 2007; Heino et al., 2017). According to the hierarchical landscape filters model of Poff (1997), species from a regional pool must pass through a series of nested filters in hierarchical order to join a local community. Until recent years, there has been a prevailing idea that unicellular organisms are ubiquitously distributed (Finlay, 2002), environmental filtering is the main mechanism structuring also diatom communities and spatial factors have only minor effects on their community structure (Finlay & Fenchel, 2004; Soininen, 2012). This has been due to the consideration that diatoms have enormous population sizes (Finlay, 2002) and are efficient passive dispersers (Kristiansen, 1996). Nevertheless, spatial factors have been shown to be important structuring elements for diatoms (Hillebrand et al., 2001; Soininen & Weckström, 2009; Heino et al., 2010), and they have been found to be important in determining diatom community structure

at continental (e.g. Potapova & Charles, 2002), regional (e.g. Heino et al., 2010) and watershed-scale (e.g. Göthe et al., 2013). However, many studies have also found that environmental conditions exceed spatial factors in importance for variation in community structure (e.g. Verleyen et al., 2009; Göthe et al., 2013). It has been suggested that the effects of spatial factors will increase with the spatial extent of the study area (Verleyen et al., 2009), and that the ratio of spatial and environmental components can be related to specific habitats (Soininen & Weckström, 2009). However, these can also be related to different ratios of ecological specialization (Pandit et al., 2009).

In this study, we examined the relative importance of environmental variables at local and catchment scale and spatial factors structuring stream diatom communities. Our aim was to study whether different species trait groups of stream diatoms show divergent responses to spatial and environmental factors and which processes are dominant in structuring a diatom metacommunity in subarctic streams. We tested whether responses to environmental and spatial variables varied between ecological guilds (i.e. high-profile, low-profile, motile and planktic guild) and between groups based on ecological specialization (i.e. generalists and specialists). Based on previous findings, we hypothesized the variation in the structure of the diatom communities as a whole to be related to both environmental and spatial variables ( $H_1$ ), but the environmental control to be more dominant ( $H_2$ ). We hypothesized weaker responses to the spatial variables due to the small study area (i.e. virtually no dispersal limitation). We also hypothesized that there would be variation in responses to environmental and spatial variables between the ecological guilds ( $H_3$ ), and that generalists and specialists would differ strongly in their responses to environmental and spatial variables ( $H_4$ ). We hypothesized that the environmental control would play a more important role in explaining the variation of specialist species ( $H_5$ ), and that the variation of generalist species would depend more on spatial factors ( $H_6$ ).

## Materials and methods

### *Study area*

This study was conducted in the Tenojoki drainage basin (centred on 70°N, 26°E). The drainage basin is located in northernmost Finland and Norway, and the main river, the River Tenojoki, flows to the Arctic Ocean (Fig. 1). The total area of the drainage basin is 16 386 km<sup>2</sup>. The study area had a mean annual temperature of -1.3 °C and a mean annual precipitation of 433 mm in the climatological normal period 1981–2010 (Pirinen et al., 2012). The study area is mainly in the subarctic deciduous birch zone and it is characterized by arctic-alpine vegetation

(Hustich, 1961). At higher altitude, barren fell tundra is typical and at low altitude there are mountain birch (*Betula pubescens* ssp. *czerepanovii*) woodlands. The study area consists mainly of Precambrian bedrock and the topography of the area is characterized by variable gently sloping fells (i.e. rounded mountains) (Mansikkaniemi, 1970). Peatlands are located mainly in the valleys between fells and they are relatively rare. The percentage of lakes is quite low (3.1 %; Korhonen & Haavanlammi, 2012) at the study area, and therefore the streams have rapid fluctuations in discharge especially in the spring season (Mansikkaniemi, 1970). The area is very sparsely populated and anthropogenic influence is minimal. Thus, headwater streams in the drainage basin range from near-pristine to pristine (Schmera et al., 2013). Stream waters are circumneutral, and nutrient levels are indicative of highly oligotrophic systems (Heino et al., 2003).

A total of 55 streams from the Finnish side of the Tenojoki drainage basin were surveyed in early June 2012. We aimed to sample all easily accessible sites that met the following criteria: (1) The length of a sampled stream must be at least 1 km. (2) The distance from the sampling site to a lake or a pond upstream had to be at least 0.5 km. (3) Only streams with permanent flow were included. (4) Large rivers (i.e. stream width >25 m, water depth >50 cm) were not included in order to get reliable and comparable samples. The size of the sampling site at each stream was approximately 50 m<sup>2</sup>. All 55 sampling sites are located in tributary streams and there are no sites in the main stem of the River Tenojoki. The distance between sampling sites furthest away from each other is 142 km.

#### *Environmental variables*

Three types of explanatory variables were used: environmental variables at local and catchment scale (Table 1) and spatial variables. We decided to divide the environmental variables into two separate groups, as stream communities are structured by the hierarchical effects of environmental variables at different scales, e.g. local environmental and catchment variables (Poff, 1997). Local variables were determined at the same time with the diatom sampling. Variables included both physical habitat and water chemistry variables. Mean width of the sampling site (m) was determined based on five cross-channel measurements. Height of the lower stream bank (area of no terrestrial vegetation; cm) and steepness of the stream bank (area of terrestrial vegetation; cm) were measured at the same locations. Height of the lower stream bank was measured from the water level to the start of terrestrial vegetation. Steepness of the upper stream bank (how many centimetres the stream bank rises in two meters' distance from the stream) was measured perpendicular to the stream. Current velocity (m s<sup>-1</sup>) and depth (cm) were measured at 30 random locations in a sampling site. Moss cover (%) and particle size classes (%) were

visually estimated at 10 1 m<sup>2</sup> plots at random locations in each sampling site. A modified Wentworth's (1922) scale of particle size classes was used: sand (0.25–2 mm), gravel (2–16 mm), pebble (16–64 mm), cobble (64–256 mm) and boulder (256–1024 mm). Based on visual estimates (%) for each plot, mean values for each site were subsequently calculated and used in all statistical analyses. Shading (%) by riparian vegetation at each sampling site was also visually estimated. Conductivity and pH were measured in the field at each sampling site using YSI device model 556 MPS (YSI Inc., Yellow Springs, OH, USA). Water samples taken during fieldwork were analyzed for iron, manganese, colour and total nitrogen following European standards. In the study area, concentration of total phosphorus is mainly below the accuracy limits of the analysis methods used (< 5 µg/l) (e.g. Heino et al., 2003). Therefore, it was not analysed in this study.

The catchment variables of each stream were calculated using ArcGIS 10.1 software (ESRI, Redlands, CA, USA), and they were based on maps acquired from the National Land Survey of Finland (Table 1). These variables consisted of drainage basin area (km<sup>2</sup>), proportion of lakes (%), length of the stream (km) and lake distance index. Lake distance index was formed using the distance to the upstream lake. This index represents the influence of the lake. There were some streams that did not have a lake upstream, and for those streams a value two times the longest distance between sampling site and lake found in the study area was given to reflect zero influence. Additionally, proportion of peatlands (%), proportion of shrub (%) and proportion of rock and cobble deposit (%) were used to mirror natural background concentrations that influence water quality, as nutrients and other chemical components are leached from drainage basin to streams to a variable degree depending on land cover type.

In addition, variables representing productivity in catchment area were used: mean and standard deviation of the NDVI (*normalized difference vegetation index*; Tucker, 1979 and Tasseled Cap greenness (Crist & Cicone, 1984). The mean and standard deviation of both variables were computed, as it has been proposed that mean values describe the average degree of productivity and standard deviation describes the variation of productivity (Parviainen et al., 2013). In addition to productivity, it has been proposed that these variables act as proxies for nutrients leaching from terrestrial areas to aquatic ecosystems (Soininen & Luoto, 2012). NDVI and greenness indexes were calculated from the Landsat 7 ETM scene (Hjort & Luoto, 2006).

Spearman's correlation test (cut-off level:  $r_s = 0.8$ ) was performed between all the environmental variables to avoid high correlations between variables. Pebble (16–64 mm), length of stream (km) and NDVI variables were excluded from further analyses based on strong correlations with other variables. There were also high correlations



between other variables, but because those variables belong to different variable groups (i.e. local or catchment), these correlations were not taken into account.

#### *Sampling and processing diatoms*

Diatom sampling and processing was carried out in accordance with the European standard (SFS-EN 13946, 2003). At each sampling site, diatoms were sampled from randomly collected cobble-sized stones from water depths of approximately 10 to 30 cm. The upper surface of the stones was scrubbed with a toothbrush and stream water, the water being pooled into one composite sample for each sampling site. In the laboratory, the diatom samples were cleaned from organic material using a strong acid solution ( $\text{HNO}_3\text{:H}_2\text{SO}_4$ ; 2:1) and mounted in a synthetic resin, Naphrax®. To determine the relative abundance of the diatom species, approximately 500 diatom valves were counted and identified to the lowest possible taxonomical level for each sample. This was done with a light microscope using differential interference contrast (1000× magnification). The identification and counting followed standard methods (SFS-EN 14407, 2005) using the Diatoms of Europe series (Lange-Bertalot, 2000, 2001, 2002, 2011) and Lange-Bertalot (2011) flora and other specialized bibliographical data when needed. Taxonomic assignments could not be made for some valves and they were omitted from analyses.

#### *Dividing diatom data matrices into different groups*

For dividing data matrices by species traits, diatoms were assigned into four ecological guilds reflecting their growth morphology. This was based on the classification made by Rimet & Bouchez (2011): low-profile, high-profile, motile and planktic guild. The low-profile guild includes species that grow very close to the substrate. These species are adapted to high current velocities and to low nutrient concentrations (Rimet & Bouchez, 2011). The high-profile guild includes species of tall stature. These species are adapted to low current velocities and high nutrient concentrations (Rimet & Bouchez, 2011). The motile guild includes species that can move actively relatively fast (Passy, 2007; Rimet & Bouchez, 2011). The planktic guild includes species that are adapted to lentic environments with morphological adaptations that enable them to resist sedimentation (e.g. *Cyclotella* spp.), and additionally nearly all filamentous diatom species (e.g. *Aulacoseira*) (Rimet & Bouchez, 2011).

Diatom species were also assigned into two groups, generalists and specialists, based on their ecological specialization. This was based on niche breadth measures determined previously by Heino & Soininen (2006) in northern Finland. The measure of niche breadth should preferably be based on a dataset different from the focal dataset in community-environment modelling. Heino & Soininen (2006) determined niche breadth that measures

amplitude in species habitat distribution using the Outlying Mean Index (OMI; Dolédec et al., 2000) analysis. This multivariate method measures the marginality of species habitat distribution, i.e. the distance between the mean habitat conditions used by a species and the mean habitat conditions across the study area (Dolédec et al., 2000). It provides two relevant niche measures, including OMI (i.e. niche position) and tolerance (i.e. niche breadth). The latter was hence used as a measure of environmental niche breadth in this study, following previous studies (Heino & Soininen, 2006; Heino & Grönroos, 2014).

The sites, in which species from all four guilds and generalist and specialist species were not found, were excluded from data analysis. Thus, there were 52 sites left for further analysis (Fig. 1). Since all the diatom species found in the study area were not included in Rimet & Bouchez's (2011) classification and Heino & Soininen's (2006) data, we formed a matrix that included all the species that belonged to any of the four guilds and another matrix that included all generalists and specialist species. Therefore, there were nine species matrices in total for further analyses (Table 2).

### *Statistical methods*

To reveal spatial patterns at multiple spatial scales and address complex patterns of spatial variation, the method of Principal Coordinates of Neighbour Matrices (PCNM; Borcard & Legendre, 2002; Borcard et al., 2004; Fig. 2) was used. The PCNM analysis creates a number of spatial variables based on Euclidean (geographical) distances between sampling sites. The Euclidean distance matrix is analysed through a principal coordinate analysis to reveal spatial relationships among sites in decreasing order of spatial scale. The result are spatial variables representing spatial structures ranging from small to large-scale across a study area. The first variables with large eigenvalues represent broad-scale variation and the last ones with small eigenvalues represent finer-scale variation (Diniz-Filho & Bini, 2005). The PCNM analysis has been used increasingly to describe spatial patterns in various organism groups (e.g. Vilmi et al., 2017), as it is effective in modelling spatial structures in biological communities at multiple scales (Dray et al., 2012). The spatial structures represented by the PCNM variables can be the result of, for example, dispersal, historical factors, or spatial autocorrelation of environmental variables or biological interaction (e.g. Dray et al., 2012). However, it is also possible that using PCNM-variables in variation partitioning overestimates the spatial component (Gilbert & Bennett, 2010; Smith & Lundholm, 2010). Spatial variables were derived from the geographical coordinates of sampling sites using the function *pcnm* in the R package PCNM (Legendre et al., 2013). In this study, only spatial variables showing positive spatial autocorrelation were employed (Borcard et al., 2011). Analyses were additionally done using east and north coordinates of the sampling

sites instead of PCNM variables, but since the PCNM variables explained more of the variation in community structure, the coordinates were excluded from the analyses.

The effects of local, catchment and spatial scale variables on diatom community structure were quantified using redundancy analysis (RDA; Rao, 1964; Fig. 2). This method evaluates how much of the variation in community structure can be explained by these variable groups. The pure and shared variations were analysed using variation partitioning through the partial redundancy analysis (pRDA; Borcard et al., 1992; Fig. 2). The aim in variation partitioning is to reveal how much of the variance in species community structure can be explained uniquely by each explanatory variable group as well as the shared variance explained by different combinations of these variable groups. Also, the unexplainable variation is revealed. With three groups of explanatory variables, the result is eight different components of variation (Fig. 3; Anderson & Gribble, 1998).

First, all species matrices were Hellinger-transformed, since the species data contained many zeros and this transformation enables the use of linear methods (Legendre & Gallagher, 2001; Fig. 2). The explanatory variables were selected for final analyses using the conservative forward selection method developed by Blanchet et al. (2008; Fig. 2). This method was used to prevent the occurrence of artificially inflated explanatory powers in models. The forward selection was carried out using function *ordiR2step* in the R package *vegan* (Oksanen et al., 2013) and it was done separately for each species matrix (i.e. low-profile guild, high-profile guild etc.). The variation partitioning was done following the protocol of Borcard et al. (1992) using the function *varpart* in the R package *vegan* (Oksanen et al., 2013). In this study, only adjusted  $R^2$  values were used, as those take into account the number of explanatory variables at each variable group and sample size (Peres-Neto et al., 2006). The significance of each testable fraction was observed using test of fraction which is based on permutation (Fig. 2). This was done by using function *anova* in the R package *vegan* (Oksanen et al., 2013). All these analyses were performed separately for each species matrices in precisely the same way.

## Results

A total of 190 diatom taxa were identified, species richness per site ranging from 19 to 55 (Table 2; Online Resource). The most common species were *Achnantidium minutissimum* (Kützing) Czarnecki s.l., *Rossithidium pusillum* (Grunow) F.E.Round & Bukhtiyarova and *Fragilaria gracilis* Øestrup. The species with the highest average abundance were *A. minutissimum* s.l., *R. pusillum* and *Fragilaria arcus* (Ehrenberg) Cleve var. *arcus*, which all belong to the low-profile guild and are generalists. From the taxa 117 species (62%) belonged to the

ecological guild classification made by Rimet & Bouchez (2011). In the sampling sites, an average of 77% of species belonged to one of the ecological guilds. In the high-profile guild, there were more species than in the other guilds. Only 57 species of the taxa (33 generalist and 24 specialist species) were found in Heino & Soininen's (2006) data. However, in the study sites, an average of 60% of species were either generalists or specialists.

Through the PCNM analysis, 15 spatial variables showing positive spatial autocorrelation were formed. The most common local variable included in the RDAs, determined by the forward selections, was moss cover (%) and the most common catchment variable was lake distance index (Table 3). Both variables, as well as the spatial variable describing broad-scale relations among sites (PCNM3), were selected for all analyses made for all species matrices. In general, the spatial variables representing the broad- and mid-scale relations among the sites were more commonly selected than the spatial variables illustrating finer-scale relations among sites.

#### *The diatom community structure*

The local and catchment environmental conditions and the spatial variables all explained the diatom community structure, yet their relative importance varied for different species matrices (Table 4). Variables describing the spatial relations among sites at broad and medium scales (PCNM 2, 3, 1, 6, 8) explained slightly more (15.1%) of the variation of the whole community structure than the other two variable groups separately. The local variables that explained the variation of the whole community structure (11.9%) were moss cover (%), proportion of boulders (%), colour (mg Pt/l) and proportion of gravel (%). The catchment variables, lake distance index, standard deviation of greenness, shrub (%) and rock and cobble deposit (%), explained almost the same amount of the variation in community structure (12.2%) than the local variables.

The variation partitioning analyses showed that for the whole community the variation in community structure was better explained by the pure spatial (4.9%) than by the pure local (2.6%) or catchment (2.5%) environmental components (Fig. 3; Table 4). The variations explained jointly by the different pairs of variable groups were approximately 4 to 5%. The shared fraction between all variable groups was 1.4%. The amount of unexplained variation was relatively large in all models, with residuals ranging from approximately 65% to 84% for different ecological guilds and from 68% to 85% for generalist and specialist species matrices.

#### *The diatom data matrices divided by ecological guilds*

Almost the same pattern as with the whole community matrix emerged when only the species found in the ecological guild classification were included (i.e. ecological guilds matrix). Here, the environmental variable groups separately also contributed less than the spatial variables to the explanation of community variation. The pure catchment component accounted for only 3.6% of the variation, while the pure spatial component explained 7.6% of the variation. However, when the different ecological guilds were analysed separately, slightly different patterns emerged. Overall, the variations in different ecological guilds were better explained by the pure effects of the local variables and the spatial variables than by the pure effects of the catchment variables. The pure local and pure spatial variables explained the same amount of the variation in the high-profile guild. The pure local component explained more of the variation in the low-profile guild and motile guild than the spatial component. In explaining the variation in the low-profile guild, the catchment component was also important. Only the variation in the planktic guild was best explained by the spatial component. The shared fractions between all variable groups ranged from approximately 0 to 4% in all guilds, but the shared fractions of the spatial variables and the catchment variables were smallest (0% or negative values to 2%). The variation in the low-profile guild was explained best, as the unexplained variation was approximately 65%.

#### *The diatom data matrices divided by ecological specialization*

Almost the same picture as with the whole community emerged when only the species found in the specialist-generalist classification were included (i.e. generalist and specialist matrix). But as with the ecological guilds, when the generalists and the specialists were analysed separately, different patterns emerged. The pure catchment component explained much more of the variation in the generalist species (10.9%) than in the specialist species (0.9%). The specialists were better explained by the pure effects of spatial variables than by the pure effects of local or catchment variables. The amount of variation that could be explained was higher for the generalists (31.9%) than for the specialists (14.7%).

## **Discussion**

In stream environments, local community structure typically portrays the effects of both environmental and spatial processes (Heino et al., 2015). Our results showed that local and catchment conditions and spatial variables all affected the organization of the subarctic diatom metacommunity with different relative importances. Our findings suggest that local conditions do not solely determine diatom metacommunity organization, but that there are also

spatially-structured patterns. Our findings also suggest that diatom communities are jointly structured by environmental filtering and spatial processes (Soininen & Weckström, 2009; Vilmi et al., 2017). These processes, however, play different roles in different species trait groups.

#### *The factors structuring entire diatom communities*

The organization of the entire diatom metacommunity was determined by spatial factors and environmental variables at local and catchment scales (supports H<sub>1</sub>). Thus, our results are consistent with earlier findings (Pan et al., 1999; see also reviews by Soininen, 2011, 2012 and references therein). However, when examining the environmental variable groups separately, our results showed that spatial variables had a relatively large effect on diatom metacommunity organization (contradicts with H<sub>2</sub>). In combination, local and catchment variables explained more variation than spatial variables alone. Previous studies have found that environmental factors exceed spatial factors in importance, and that stream communities are mostly under abiotic control (Verleyen et al., 2009; Göthe et al., 2013). Our findings are in contrast with many specific studies that suggest that diatom community structures primarily reflect variation in local conditions (De Bie et al., 2012; Gottschalk & Kahlert, 2012). Strong spatial patterns have previously been found mainly in large-scale studies, as in Heino et al.'s (2010) study concerning boreal stream diatom communities, or in highly connected environments, as in Vilmi et al.'s (2017) study in a boreal lake system. Indeed, these differences in findings may be due to different spatial scales (Mykrä et al., 2007) and environmental variables examined, but also to different ratios of ecological guilds (Göthe et al., 2013; Vilmi et al., 2017) and ecological specialization (Pandit et al., 2009).

#### *The factors structuring ecological guilds*

Our results showed that there was variation in responses to environmental and spatial variables between the ecological guilds (supports H<sub>3</sub>). Overall, the variations in different ecological guilds were better explained by the local and spatial variables than by catchment variables. Our findings suggest that the high- and low-profile guilds are simultaneously structured by environmental filtering and spatial processes in subarctic streams. However, environmental filtering plays a more important role for the motile guild, and spatial-related processes are important for planktic species. The planktic guild has shown clear spatial patterns in other studies as well (e.g. Vilmi et al., 2017). In boreal streams (Göthe et al., 2013) and lakes (Vilmi et al., 2017), diatom guilds have also been structured by various metacommunity processes. Göthe et al. (2013) suggested that the dissimilar findings between guilds could be due to diatoms' traits related to dispersal capacity. According to Algarte et al. (2014),

firmly attached algae (i.e. low-profile guild species) show clear spatial patterns, as they resist high current velocities (Passy, 2007). Thus, they have lower dispersal rates. In our study, this was not the case, as the local environmental component explained best the variation in the low-profile guild. It has been also suggested that the degree of attachment and the mobility of micro-organisms can affect the extent of dispersal (Vilmi et al., 2017). This can partly explain the importance of spatial-related processes to planktic guild species in our study. Unfortunately, dispersal capacities of diatom species and what traits determine them—at least in terms of long-distance dispersal—is a subject that has not been studied much (Kristiansen, 1996; Vyverman et al., 2007; Casteleyn et al., 2010; Souffreau et al. 2013; Rimet et al., 2014). However, the use of guild division can give us some indirect indications of dispersal processes.

#### *The factors affecting different groups of ecological specialization*

Our results showed that generalists and specialists differ strongly in responses to environmental and spatial variables (supports H<sub>4</sub>; Pandit et al., 2009; Székely & Langenheder, 2014). We thought that generalists would be structured by spatial-related processes because they can tolerate a wide range of environmental conditions (Devictor et al., 2010). However, the variation in the generalist species was explained mostly by catchment environmental factors (contradicts with H<sub>6</sub>). According to the hierarchical environmental filtering model (Poff, 1997), regional processes determine the species reaching the local habitat. Thus, it is possible that regional processes are limiting factors to generalist species. Our results also indicated that spatial processes are important to specialist species (contradict H<sub>5</sub>). Dispersal can be more challenging to specialist species because there are fewer suitable environments for them (Kolasa & Romanuk, 2005). However, it is unlikely that dispersal limitation would explain these spatial patterns due to the relatively small spatial extent of our study area and the fact that this study was conducted within one drainage basin (see Mouquet & Loreau, 2003; Leibold et al., 2004; Heino et al., 2017).

Our results are slightly inconsistent with previous studies (e.g. Pandit et al., 2009). With rock pool invertebrates, habitat generalists respond mainly to spatial factors and habitat specialists mostly to environmental factors (Pandit et al., 2009). On the other hand, community composition of generalist bacteria was best explained by environmental factors (Székely & Langenheder, 2014). In addition, for dragonflies, dispersal restricted the distributions of habitat specialist species (McCauley, 2007). In Alahuhta et al.'s (2014) study, the community compositions of both common and rare macrophyte species were explained by environmental factors, suggesting environmental filtering to be more dominant regardless of the degree of rarity.

In our study, the amount of explained variation was much higher for the generalists than for the specialists. This is not surprising, as specialist species have a narrower niche breadth, and environmental factors can affect different specialist species in different ways (Pandit et al., 2009). Overall, some species can be strongly specialized or clearly generalists, but generally, species are something in between these extreme ends (Heino & Soininen, 2006; Pandit et al., 2009). Thus, the generalist and specialist division in our study is rather coarse. However, our results suggest that even this coarse division can be useful when studying the effects of ecological specialization on community structure.

#### *Spatial processes and scale dependency*

Our results showed that spatial variables had a much larger effect on diatom metacommunity organization than we thought based on the relatively small spatial extent of our study area (Verleyen et al., 2009; Bennett et al., 2010). However, Astorga et al. (2012) have found that diatom communities are spatially structured in very similar environments at small scale (<200 km) but not at larger spatial extents. In studies concerning microbial communities, spatial patterns have been found at the small spatial scale in systems of high connectivity (Lear et al., 2014; Vilmi et al., 2016; Vilmi et al., 2017). Connectivity probably can also play a role in stream diatom metacommunities. Historical factors are important in explaining geographical patterns found in diatom genus richness at regional to global scales, indicating the vital roles of dispersal limitation in structuring diatom communities (Vyverman et al., 2007). Thus, as the spatial variables used in this study can portray also the historical factors and dispersal (Dray et al., 2012), this could explain the importance of these variables also in our study, although the scale in our study is much smaller. However, spatial structures found in small spatial extent and within a region (i.e. Tenojoki drainage basin) are usually mainly related to homogenizing effects rather than dispersal limitations (Mouquet & Loreau, 2003; Leibold et al., 2004; Heino et al., 2017), even though both can produce spatial patterns (Ng et al., 2009). These homogenizing effects can take place via mass-effects (Mouquet & Loreau, 2003). In the Tenojoki drainage basin, diatom communities seem to be structured by processes active at multiple spatial scales, as they have been in comparable studies (Göthe et al., 2013; Vilmi et al., 2016; Vilmi et al., 2017). However, interpretation of spatial variables is always dependent on the size and connectivity of the study system (Dray et al., 2012).

#### *Concluding remarks*



The results of this study should be interpreted with caution, as the amounts of unexplained variation were relatively high. This was partly due to the statistical methods used (adjusted coefficient of determination; Peres-Neto et al., 2006), and low amount of explained variation is common in these kind of studies (e.g. Pandit et al., 2009; Algarte et al., 2014). Moreover, it is possible that some important explanatory variables are missing from the analysis (e.g. Algarte et al., 2014). For example, this study did not include biotic interaction, e.g. grazing. However, previous studies have shown that grazing has no apparent effects, at least on the structure of diatom guilds (e.g. Göthe et al., 2013; Vilmi et al., 2017). Yet, biotic and trophic interactions would be an interesting addition to the study of northern, nutrient-poor environments. According to Berthon et al. (2011), grazing pressure may be higher in nutrient-poor rivers than in nutrient-rich rivers because biofilms are rare. However, a more likely reason for the low amounts of explained variations is the occurrence of stochastic processes (Vellend et al., 2014), as biological communities are formed through very complex processes and interactions. The guild and ecological specialization information were not available for all species and this can have implications on results. However, we believe that our results are representative, because the reduced overall guild and ecological specialization matrices showed patterns similar to those of the entire community matrix.

Our findings suggested that dividing the whole community into different groups by species traits indeed increases understanding of metacommunity organization. Our study showed that diatom communities in subarctic streams are a result of both environmental filtering and spatial-related processes. Future studies should focus on measuring grazing pressure, especially in nutrient-poor subarctic streams, and dispersal rates of diatom species to acquire more reliable knowledge of the processes structuring diatom communities. Focusing on these biological processes would, however, necessitate experimental approaches, which may be complicated at spatial extents comprising entire drainage basins. Hence, large-scale observational studies offer necessary background information for guiding more detailed experimental work and provide important information for biodiversity assessment research.

## References

- Alahuhta, J., L. B. Johnson, J. Olker & J. Heino, 2014. Species sorting determines variation in the community composition of common and rare macrophytes at various spatial extents. *Ecological Complexity* 20: 61–68.
- Algarte, V. M., L. Rodrigues, V. L. Landeiro, T. Siqueira & L. M. Bini, 2014. Variance partitioning of deconstructed periphyton communities: Does the use of biological traits matter? *Hydrobiologia* 722: 279–290.
- Anderson, M. J. & N. A. Gribble, 1998. Partitioning the variation among spatial, temporal and environmental components in a multivariate data set. *Australian Journal of Ecology* 23: 158–167.
- Astorga, A., J. Oksanen, M. Luoto, J. Soininen, R. Virtanen & T. Muotka, 2012. Distance decay of similarity in freshwater communities: Do macro- and microorganisms follow the same rules? *Global Ecology and Biogeography* 21: 365–375.
- Bennett, J. R., B. F. Cumming, B. K. Ginn & J. P. Smol, 2010. Broad-scale environmental response and niche conservatism in lacustrine diatom communities. *Global Ecology and Biogeography* 19: 724–732.
- Berthon, V., A. Bouchez & F. Rimet, 2011. Using diatom lifeforms and ecological guilds to assess organic pollution and trophic level in rivers: A case study of rivers in southeaster France. *Hydrobiologia* 673: 259–271.
- Blanchet, F. G., P. Legendre & D. Borcard, 2008. Forward selection of explanatory variables. *Ecology* 89: 2623–2632.
- Borcard, D., F. Gillet & P. Legendre, 2011. *Numerical Ecology with R*. Springer, New York.
- Borcard, D. & P. Legendre, 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological Modelling* 153: 51–68.
- Borcard, D., P. Legendre, C. Avois-Jacquet & H. Tuomisto, 2004. Dissecting the spatial structure of ecological data at multiple scales. *Ecology* 85: 1826–1832.
- Borcard, D., P. Legendre & P. Drapeau, 1992. Partialling out the spatial component of ecological variation. *Ecology* 73: 1045–1055.
- Cadotte, M. W. & C. M. Tucker, 2017. Should environmental filtering be abandoned? *Trends in Ecology and Evolution* 32: 429–437.

478 Casteleyn, G., F. Leliaert, T. Backeljau, A.-E. Debeer, Y. Kotaki, L. Rhodes et al., 2010. Limits to gene flow in  
479 a cosmopolitan marine planktonic diatom. *Proceedings of the National Academy of Sciences* 107:  
480 12952–12957.

481 Crist E. P. & R. C. Cicone, 1984. A physically-based transformation of thematic mapper data – the TM tasseled  
482 cap. *IEEE Transactions on Geoscience and Remote Sensing* 22: 256–263.

483 De Bie, T., L. De Meester, L. Brendonck, K. Martens, B. Goddeeris, D. Ercken et al., 2012. Body size and  
484 dispersal mode as key traits determining metacommunity structure of aquatic organisms. *Ecology Letters*  
485 15: 740–747.

486 Devictor, V., J. Clavel, R. Julliard, S. Lavergne, D. Mouillot, W. Thuiller et al., 2010. Defining and measuring  
487 ecological specialization. *Journal of Applied Ecology* 47: 15–25.

488 Devictor, V., R. Julliard & F. Jiguet, 2008. Distribution of specialist and generalist species along spatial  
489 gradients of habitat disturbance and fragmentation. *Oikos* 117: 507–514.

490 Diniz-Filho, J. A. F. & L. M. Bini, 2005. Modelling geographical patterns in species richness using eigenvector-  
491 based spatial filters. *Global Ecology and Biogeography* 14: 177–185.

492 Dolédec, S., D. Chessel & C. Gimaret-Carpentier, 2000. Niche separation in community analysis: A new  
493 method. *Ecology* 81: 2914–2927.

494 Dray, S., R. Péliissier, P. Couteron, M. J. Fortin, P. Legendre, P. R. Peres-Neto et al., 2012. Community ecology  
495 in the age of multivariate multiscale spatial analysis. *Ecological Monographs* 82: 257–275.

496 Finlay, B. J., 2002. Global dispersal of free-living microbial eukaryote species. *Science* 296: 1061–1063.

497 Finlay, B. J. & T. Fenchel, 2004. Cosmopolitan metapopulations of free-living microbial eukaryotes. *Protist*  
498 155: 237–244.

499 Gilbert, B. & J. R. Bennett, 2010. Partitioning variation in ecological communities: Do the numbers add up?  
500 *Journal of Applied Ecology* 47: 1071–1082.

501 Gottschalk, S. & M. Kahlert, 2012. Shifts in taxonomical and guild composition of littoral diatom assemblages  
502 along environmental gradients. *Hydrobiologia* 694: 41–56.

503 Grönroos, M., J. Heino, T. Siqueira, V. L. Landeiro, J. Kotanen & L. M. Bini, 2013. Metacommunity structuring  
504 in stream networks: Roles of dispersal mode, distance type, and regional environmental context. *Ecology*  
505 and Evolution 3: 4473–4487.

506 Göthe, E., D. G. Angeler, S. Gottschalk, S. Löfgren & L. Sandin, 2013. The influence of environmental, biotic  
507 and spatial factors on diatom metacommunity structure in swedish headwater streams. *PLoS ONE* 8: 1–9.

508 Hájek, M., J. Roleček, K. Cottenie, K. Kintrová, M. Horsák, A. Pouličková et al., 2011. Environmental and  
 509 spatial controls of biotic assemblages in a discrete semi-terrestrial habitat: Comparison of organisms with  
 510 different dispersal abilities sampled in the same plots. *Journal of Biogeography* 38: 1683–1693.

511 Heino, J., L. M. Bini, S. M. Karjalainen, H. Mykrä, J. Soininen, L. C. G. Vieira & J. A. F. Dini-Filho, 2010.  
 512 Geographical patterns of micro-organismal community structure: Are diatoms ubiquitously distributed  
 513 across boreal streams? *Oikos* 119: 129–137.

514 Heino, J. & M. Grönroos, 2014. Untangling the relationships among regional occupancy, species traits and niche  
 515 characteristics in stream invertebrates. *Ecology and Evolution* 4: 1931–1942.

516 Heino, J., A. S. Melo, T. Siqueira, J. Soininen, S. Valanko & L. M. Bini, 2015. Metacommunity organisation,  
 517 spatial extent and dispersal in aquatic systems: Patterns, processes and prospects. *Freshwater Biology* 60:  
 518 845–869.

519 Heino, J., T. Muotka & R. Paavola, 2003. Determinants of macroinvertebrate diversity in headwater streams:  
 520 Regional and local influences. *Journal of Animal Ecology* 72: 425–434.

521 Heino, J. & J. Soininen, 2006. Regional occupancy in unicellular eukaryotes: A reflection of niche breadth,  
 522 habitat availability or size-related dispersal capacity? *Freshwater Biology* 51: 672–685.

523 Heino, J., J. Soininen, J. Alahuhta, J. Lappalainen & R. Virtanen, 2017. Metacommunity ecology meets  
 524 biogeography: Effects of geographical region, spatial dynamics and environmental filtering on  
 525 community structure in aquatic organisms. *Oecologia* 183: 121–137.

526 Hillebrand, H., F. Watermann, R. K. Arez & U. G. Berninger, 2001. Differences in species richness patterns  
 527 between unicellular and multicellular organisms. *Oecologia* 126: 114–124.

528 Hjort, J. & M. Luoto, 2006. Modelling patterned ground distribution in Finnish Lapland: An integration of  
 529 topographical, ground and remote sensing information. *Geografiska Annaler Series A-physical*  
 530 *Geography* 88A: 19–29.

531 Holyoak, M., M. A. Leibold, N. Mouquet, R. D. Holt & M. F. Hoops, 2005. Metacommunities: A framework  
 532 for large-scale community ecology. In Holyoak M. et al. (eds), *Metacommunities: Spatial Dynamics and*  
 533 *Ecological Communities*. The University of Chicago Press, Chicago: 1–31.

534 Hustich, I., 1961. Plant geographical regions. In Somme A. (ed), *A Geography of Norden*. Heinemann, Oslo:  
 535 54–62.

536 Jacobson, B. & P. R. Peres-Neto, 2010. Quantifying and disentangling dispersal in metacommunities: How  
 537 close have we come? How far is there to go? *Landscape Ecology* 25: 495–507.

538 Kolasa, J. & T. N. Romanuk, 2005. Assembly of unequals in the unequal world of a rock pool metacommunity.  
 539 In Holyoak M. et al. (eds), *Metacommunities: Spatial Dynamics and Ecological Communities*. The  
 540 University of Chicago press, Chicago: 212–232.

541 Korhonen, J. & E. Haavanlammi, 2012. Hydrological Yearbook 2006–2010. *The Finnish Environment* 8/2012.

542 Kristiansen, J., 1996. Dispersal of freshwater algae – a review. *Hydrobiologia* 336: 151–157.

543 Lange-Bertalot, H. (ed), 2000–2011. *Diatoms of Europe: Diatoms of the European Inland Waters and*  
 544 *Comparable Habitats*. Volumes 1–6. Ruggell: A.R.G. Gantner Verlag K.G.

545 Lange-Bertalot, H. (ed), 2011. *Diatomeen im Süßwasser - Benthos von Mitteleuropa*. A. R. G. Gantner Verlag  
 546 K. G., Ruggell.

547 Lear, G., J. Bellamy, B. S. Case, J. E. Lee & H. L. Buckley, 2014. Fine-scale spatial patterns in bacterial  
 548 community composition and function within freshwater ponds. *The ISME Journal* 8: 1715–26.

549 Legendre, P. & D. E. Gallagher, 2001. Ecologically meaningful transformations for ordination of species data.  
 550 *Oecologia* 129: 271–280.

551 Legendre, P., D. Borcard, F. G. Blanchet & S. Dray, 2013. PCNM: MEM spatial eigenfunction and principal  
 552 coordinate analyses. R package version 2.1-2. Available at: [http://r-forge.r-project.org/R/?group\\_id=195](http://r-forge.r-project.org/R/?group_id=195).

553 Leibold, M. A., M. Holyoak, N. Mouquet, P. Amarasekare, J. M. Chase, M. F. Hoopes et al., 2004. The  
 554 metacommunity concept: A framework for multi-scale community ecology. *Ecology Letters* 7: 601–613.

555 Lindström, E. S. & S. Langenheder, 2012. Local and regional factors influencing bacterial community assembly.  
 556 *Environmental Microbiology Reports* 4: 1–9.

557 Mansikkaniemi, H., 1970. Deposits of sorted material in the Inarijoki-Tana river valley in Lapland. *Reports of*  
 558 *Kevo Subarctic Research Station* 6: 1–63.

559 McCauley, S. J., 2007. The role of local and regional processes in structuring larval dragonfly distributions  
 560 across habitat gradients. *Oikos* 116: 121–133.

561 Mouquet, N. & M. Loreau, 2003. Community patterns in source-sink metacommunities. *American Naturalist*  
 562 162: 544–557.

563 Mykrä, H., J. Heino & T. Muotka, 2007. Scale-related patterns in the spatial and environmental components of  
 564 stream macroinvertebrate assemblage variation. *Global Ecology and Biogeography* 16: 149–159.

565 National Board of Waters and the Environment, 1981. *Vesihallinnon analyysimenetelmät*. Tiedotus 213.  
 566 Vesihallitus, Helsinki.

567 Ng, I. S. Y., C. M. Carr & K. Cottenie, 2009. Hierarchical zooplankton metacommunities: distinguishing  
568 between high and limiting dispersal mechanisms. *Hydrobiologia* 619: 133–143.

569 Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O’Hara et al., 2013. Vegan:  
570 Community Ecology package. R package version 2.0-7. Available at: [https://cran.r-](https://cran.r-project.org/web/packages/vegan/index.html)  
571 [project.org/web/packages/vegan/index.html](https://cran.r-project.org/web/packages/vegan/index.html)

572 Pan, Y., R. J. Stevenson, B. H. Hill, P. R. Kaufmann & A. T. Herlihy, 1999. Spatial Patterns and Ecological  
573 Determinants of Benthic Algal Assemblages in Mid-Atlantic Streams, USA. *Journal of Phycology* 35:  
574 460–468.

575 Pandit, S. N., J. Kolasa & K. Cottenie, 2009. Contrasts between habitat generalists and specialists: An empirical  
576 extension to the basic metacommunity framework. *Ecology* 90: 2253–2262.

577 Parviainen, M., N. E. Zimmermann, R. K. Heikkinen & M. Luoto, 2013. Using unclassified continuous remote  
578 sensing data to improve distribution models of red-listed plant species. *Biodiversity and Conservation* 22:  
579 1731–1754.

580 Passy, S. I., 2007. Diatom ecological guilds display distinct and predictable behavior along nutrient and  
581 disturbance gradients in running waters. *Aquatic Botany* 86: 171–178.

582 Peres-Neto, P. R., P. Legendre, S. Dray & D. Borcard, 2006. Variation partitioning of species data matrices:  
583 Estimation and comparison of fractions. *Ecology* 87: 2614–2625.

584 Pirinen, P., H. Simola, J. Aalto, J.-P. Kaukoranta, P. Karlsson & R. Ruuhela, 2012. Climatological statistics of  
585 Finland 1981-2010. Reports 2012: 1. Finnish Meteorological Institute, Helsinki.

586 Poff, N. L., 1997. Landscape filters and species traits: Towards mechanistic understanding and prediction in  
587 stream ecology. *Journal of the North American Benthological Society* 16: 391–409.

588 Potapova, M. G. & D. F. Charles, 2002. Benthic diatoms in USA rivers: Distributions along spatial and  
589 environmental gradients. *Journal of Biogeography* 29: 167–187.

590 Rao, C. R., 1964. The use and interpretation of principal component analysis in applied research. *Sankhyā: The*  
591 *Indian Journal of Statistics, Series A* 26: 329–358.

592 Rimet, F. & A. Bouchez, 2012. Life-forms, cell-sizes and ecological guilds of diatoms in European rivers.  
593 *Knowledge and Management of Aquatic Ecosystems* 406: 1–14.

594 Rimet, F., R. Trobajo, D. G. Mann, L. Kermarrec, A. Franc, I. Domaizon & A. Bouchez, 2014. When is  
595 sampling complete? The effects of geographical range and marker choice on perceived diversity in  
596 *Nitzschia palea* (Bacillariophyta). *Protist* 165: 245–259.

597 Schmera, D., T. Erös & J. Heino, 2013. Habitat filtering determines spatial variation of macroinvertebrate  
 598 community traits in northern headwater streams. *Community Ecology* 14: 77–88.  
 599 SFS-EN 13946, 2003. Veden laatu. Jokivesien piilevien näytteenotto ja esikäsittely. Suomen standardoimisliitto  
 600 SFS ry, Helsinki.  
 601 SFS-EN 14407, 2005. Water quality. Guidance standard for the identification, enumeration and interpretation of  
 602 benthic diatom samples from running waters. Suomen standardisoimisliitto SFS ry, Helsinki.  
 603 Smith, T. W. & J. T. Lundholm, 2010. Variation partitioning as a tool to distinguish between niche and neutral  
 604 processes. *Ecography* 33: 648–655.  
 605 Soininen, J., 2011. Environmental and spatial control of freshwater diatoms – a review. *Diatom Research* 22:  
 606 473–490.  
 607 Soininen, J., 2012. Macroecology of unicellular organisms – patterns and processes. *Environmental*  
 608 *Microbiology Reports* 4: 10–22.  
 609 Soininen, J. & M. Luoto, 2012. Is catchment productivity a useful predictor of taxa richness in lake plankton  
 610 communities? *Ecological Applications* 22: 624–633.  
 611 Soininen, J. & J. Weckström, 2009. Diatom community structure along environmental and spatial gradients in  
 612 lakes and streams. *Fundamental and Applied Limnology* 174: 205–213.  
 613 Souffreau, C., P. Vanormelingen, K. Sabbe & W. Vyverman, 2013. Tolerance of resting cells of freshwater and  
 614 terrestrial benthic diatoms to experimental desiccation and freezing is habitat-dependent. *Phycologia* 52:  
 615 246–255.  
 616 Székely, A. J. & S. Langenheder, 2014. The importance of species sorting differs between habitat generalists  
 617 and specialists in bacterial communities. *FEMS Microbiology Ecology* 87: 102–112.  
 618 Tucker, C. J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote*  
 619 *sensing of environment* 8: 127–150.  
 620 Van der Gucht, K., K. Cottenie, K. Muylaert, N. Vloemans, S. Cousin, S. Declerck et al., 2007. The power of  
 621 species sorting: Local factors drive bacterial community composition over a wide range of spatial scales.  
 622 *Proceedings of the National Academy of Sciences of the United States of America* 104: 20404–20409.  
 623 Verleyen, E., W. Vyverman, M. Sterken, D. A. Hodgson, A. De Wever, S. Juggins et al., 2009. The importance  
 624 of dispersal related and local factors in shaping the taxonomic structure of diatom metacommunities.  
 625 *Oikos* 118: 1239–1249.

626 Vilmi, A., S. M. Karjalainen, S. Hellsten & J. Heino, 2016. Bioassessment in a metacommunity context: Are  
627 diatom communities structured solely by species sorting? *Ecological Indicators* 62: 86–94.

628 Vilmi, A., K. T. Tolonen, S. M. Karjalainen & J. Heino, 2017. Metacommunity structuring in a highly-  
629 connected aquatic system: Effects of dispersal, abiotic environment and grazing pressure on microalgal  
630 guilds. *Hydrobiologia* 790: 125–140.

631 Vyverman, W., E. Verleyen, K. Sabbe, K. Vanhoutte, M. Sterken, D. A. Hodgson et al., 2007. Historical  
632 processes constrain patterns in global diatom diversity. *Ecology* 88: 1924–1931.

633 Wentworth, C. K., 1922. A scale of grade and class terms for clastic sediments. *Journal of Geology* 30: 377–  
634 392.

635 Wilson, D. S., 1992. Complex interactions in metacommunities, with implications for biodiversity and higher  
636 levels of selection. *Ecology* 73: 1984–2000.



637 **Tables and figures**

638 **Table 1** Summary of local and catchment variables across the study sites in the River Tenojoki drainage basin.

639 N = 52 streams.

Variable	Min	Max	Mean	SD
<b>Local scale</b>				
Total nitrogen (µg/l)	62	260	132.08	43.8
Color (mg Pt/l)	10	50	27.40	9.62
Iron (µg/l)	8	160	69.06	41.06
Manganese (µg/l)	1	5.5	2.01	1.39
pH	6.58	7.51	6.87	0.17
Conductivity (µS/cm)	11	27	17.48	3.96
Particle size (%)				
Sand (0.25–2 mm)	0	24.5	0.88	3.48
Gravel (2–16 mm)	0	12	2.62	2.95
Pebble (16–64 mm)	0	45.67	14.40	11.17
Cobble (64–256 mm)	1	52	24.51	11.82
Boulder (256–1024 mm)	14	99	57.60	21.08
Moss cover (%)	0.3	75	17.76	20.21
Current velocity (m/s)	0.28	0.89	0.57	0.13
Depth (cm)	14.6	34.47	24.27	4.62
Mean width (m)	1.2	22	5.91	4.22
Height of the lower stream bank (cm)	0.0	117.9	32.00	24.6
Steepness (cm)	0.5	108	36.74	23.41
Shading (%)	0	100	41.46	33.58
<b>Catchment scale</b>				
Drainage basin area (km <sup>2</sup> )	1.55	135.74	24.89	29.95
Proportion of lakes (%)	0	11	1.27	2.27
Lake distance index	1.14	52.51	30.76	24.09
Length of the stream (km)	1.39	28.97	9.28	6.8
Peatlands (%)	1.17	39.78	12.70	8.13
Shrub (%)	0	93.87	45.18	31.55
Rock and cobble deposit (%)	0	26.88	2.89	4.23
NDVI, mean	-0.03	0.57	0.26	0.14
NDVI, standard deviation	0.1	0.33	0.21	0.05
Greenness, mean	0.09	0.19	0.14	0.02
Greenness, standard deviation	0.02	0.05	0.03	0.01

640 Minimum (min), maximum (max) and mean (mean) values and standard deviation (SD).

641 **Table 2** Total number of diatom species, and minimum (min), maximum (max), mean (mean) and standard  
642 deviation (SD) of local number of species in different species matrices.

Species matrix	Number of species	Min	Max	Mean	SD
All taxa	190	19	55	32.5	8.18
Ecological guilds	117	14	40	24.98	5.93
High-profile guild	46	3	17	8.75	3.03
Low-profile guild	33	5	15	9.88	2.53
Motile guild	27	1	9	3.42	1.96
Planktic guild	11	1	5	2.92	0.97
Generalists and specialists	57	10	31	19.21	4.37
Generalist	33	7	23	13.96	3.37
Specialist	24	1	10	5.25	2.25

643

**Table 3** The selected variables according to the forward selection procedure and their rank order.

	<b>Local</b>	<b>Catchment</b>	<b>Spatial</b>
<b>All taxa</b>	Moss cover	Lake distance index	PCNM2
	Boulder	Greenness, standard deviation	PCNM3
	Color	Shrub	PCNM1
	Gravel	Rock and cobble deposit	PCNM6
			PCNM8
<b>Ecological guilds</b>	Moss cover	Lake distance index	PCNM2
	Gravel	Greenness, standard deviation	PCNM3
	Manganese	Rock and cobble deposit	PCNM8
	Iron		PCNM1
	Current velocity		PCNM9
	Shading		PCNM6
<b>High-profile</b>	Moss cover	Lake distance index	PCNM3
	Conductivity	Greenness, mean	PCNM13
	Manganese		
	Color		
<b>Low-profile</b>	Moss cover	Lake distance index	PCNM2
	Gravel	Greenness, standard deviation	PCNM3
	Current velocity	Shrub	PCNM8
	Shading		
<b>Motile</b>	Boulder	Lake distance index	PCNM3
	Moss cover	Shrub	PCNM15
	Iron		PCNM11
	Current velocity		
<b>Planktic</b>	Moss cover	Lake distance index	PCNM3
	Conductivity	Peatlands	PCNM9
	Boulder		PCNM2
<b>Generalists and specialists</b>	Moss cover	Lake distance index	PCNM2
	Gravel	Greenness, standard deviation	PCNM3
	Current velocity	Rock and cobble deposit	PCNM8
	Manganese		PCNM6
	Iron		PCNM1
			PCNM9
<b>Generalist</b>	Moss cover	Lake distance index	PCNM2
	Gravel	Greenness, standard deviation	PCNM8
	Current velocity	Rock and cobble deposit	PCNM3
<b>Specialist</b>	Moss cover	Lake distance index	PCNM3
	Manganese	Drainage basin area	PCNM13
			PCNM9

646 **Table 4** A summary of variation partitioning results. The significance of each fraction explained is indicated in the table (\*p<0.05, \*\*p<0.01).

	Generalists																	
	All taxa		Ecological		High-profile		Low-profile		Planktic		plus		Generalists		Specialists			
			guilds		guild		guild		Motile guild		guild		specialists					
	Df	Adj. R <sup>2</sup>	Df	Adj. R <sup>2</sup>	Df	Adj. R <sup>2</sup>	Df	Adj. R <sup>2</sup>	Df	Adj. R <sup>2</sup>	Df	Adj. R <sup>2</sup>	Df	Adj. R <sup>2</sup>	Df	Adj. R <sup>2</sup>	Df	Adj. R <sup>2</sup>
Local [a+d+f+g]	4	0.119	6	0.161	4	0.099	4	0.189	4	0.095	3	0.136	5	0.156	3	0.143	2	0.063
Catchment [b+d+e+g]	4	0.121	3	0.119	2	0.055	3	0.149	2	0.055	2	0.109	3	0.128	3	0.154	2	0.043
Spatial [c+e+f+g]	5	0.151	6	0.159	2	0.062	3	0.155	3	0.076	3	0.143	6	0.158	3	0.120	3	0.096
[a+b+d+e+f+g]	8	0.190	9	0.212	6	0.121	7	0.279	6	0.132	5	0.177	8	0.217	6	0.249	4	0.085
[a+c+d+e+f+g]	9	0.213	12	0.252	6	0.139	7	0.281	7	0.134	6	0.240	11	0.249	6	0.210	5	0.138
[b+c+d+e+f+g]	9	0.212	9	0.233	4	0.118	6	0.252	5	0.115	5	0.206	9	0.243	6	0.260	5	0.120
[a+b+c+d+e+f+g]	13	0.238	15	0.288	8	0.161	10	0.351	9	0.159	8	0.266	14	0.295	9	0.319	7	0.147
Individual fractions																		
[a] Pure local	4	0.026*	6	0.055*	4	0.043*	4	0.099*	4	0.044*	3	0.060*	5	0.052*	3	0.058*	2	0.027*
[b] Pure catchment	4	0.025*	3	0.036*	2	0.023*	3	0.070*	2	0.025*	2	0.025*	3	0.046*	3	0.109*	2	0.009
[c] Pure spatial	5	0.049*	6	0.076*	2	0.040*	3	0.072*	3	0.026	3	0.088*	6	0.078*	3	0.070*	3	0.062*
[d] Local + catchment	0	0.036	0	0.037	0	0.034	0	0.028	0	0.014	0	0.037	0	0.039	0	0.032	0	0.015
[e] Catchment + spatial	0	0.046	0	0.015	0	0	0	0.020	0	0.013	0	0.016	0	0.015	0	-0.003	0	0.014
[f] Local + spatial	0	0.042	0	0.037	0	0.023	0	0.032	0	0.034	0	0.009	0	0.037	0	0.036	0	0.015
[g] Shared	0	0.014	0	0.031	0	0.001	0	0.031	0	0.003	0	0.030	0	0.028	0	0.016	0	0.005
Residuals [h]		0.762		0.712		0.839		0.649		0.841		0.735		0.705		0.681		0.853

647

648 **Fig. 1** Map showing the location of the Tenojoki drainage basin, the study sites and the catchments of those sites  
649 (green). Only the streams from the Finnish side of the Tenojoki drainage basin are presented, with the exception of the  
650 main stem of River Tenojoki and the most north-eastern part of the map. Note that all 52 study sites are located in  
651 tributary streams and there are no sites in the main stem of the River Tenojoki. Only sites included in the data analyses  
652 are visible on the map

653

654

655 **Fig. 2** A schematic diagram showing the methodology used. The analyses were done separately for each species data  
656 matrix

657

658

659 **Fig. 3** Venn-diagrams showing the fractions of diatom community structure explained by the local variables (Local), the  
660 catchment variables (Catchment) and spatial variables (Spatial). All fractions are based on adjusted  $R^2$  values shown as  
661 percentages of total variation. Values  $<0$  are not shown. A = all taxa, B = ecological guilds, C = high-profile guild, D =  
662 low-profile guild, E = motile guild, F = planktic guild, G = generalist and specialist, H = generalist, I = specialist